RELATIONSHIPS OF LABOR COSTS TO SELECTED VARIABLES IN FLUE-CURED TOBACCO PRODUCTION

Garnett L. Bradford*

Since 1960, data from "typical" enterprise budgets have been used extensively in several studies dealing with a wide variety of problems [5, 6, 10]. Budgets were compiled in cost-of-production studies by Green [4], Pierce and Williams [9], Coutu and Mangum [2] and the North Carolina Agricultural Extension Service [8]. Without surprise, extensive use of these budgets has focused attention on their limitations.

More extensive and precise field measurements were needed for several cost items, particularly labor costs. In addition, certain types of cost-input or cost-output relationships were implicit in such constant-type coefficients—relationships which may or may not exist in reality. It may be true, for example, that X hours of priming labor are required to harvest Y pounds of tobacco, but it might require α + 1.8X hours to harvest 2Y pounds.

The primary objectives of this discussion are (1) to discuss procedures for estimating the relationships between tobacco labor costs and selected production variables, and (2) to report some test results of hypotheses about the nature of the cost-input or cost-output functional relationships in the conventional production of flue-cured tobacco. A secondary objective is to illustrate adaptations of existing experimental design and statistical techniques developed, in the study, for use in cost-of-production studies; specifically, the use of mixed regression or covariance estimation models, as a method of meeting the primary objectives, is illustrated.

DELINEATION OF HYPOTHESES

Unit cost measurements, like those quantified by Bradford and Nelson [1], can be used to estimate per acre costs for different yield levels or production practices. However, such a procedure implicity assumes that priming cost per acre is a linear function of pounds (weight) with a zero intercept. It also assumes that the relationship does not vary between experiment locations and years, or between leaf positions on the tobacco plant.

These assumptions were delineated into sequences of hypotheses. Each sequence of hypotheses was tested for all labor operations commonly included in a flue-cured tobacco budget. For example, the following hypotheses were tested for priming labor: (1) Priming labor is a linear function of harvested leaves. (2) The function has a zero intercept. (3) The same function applies to all years and locations (of the study) and to all stalk positions of the tobacco plant.

Data were obtained primarily from measurements made in controlled experiments and were analyzed by fitting mixed regression models.

MEASUREMENT PROCEDURES

To obtain labor time measurements corresponding to relatively diverse levels of inputs, experiments were conducted in 1963 through 1965 at four farm locations in North Carolina—three farms per year. In each experiment were two blocked replications of each of three basic treatments. Each treatment consisted of combinations of fertilizer, sucker control materials, plants per acre and topping heights. There were 112, 151, and 190 thousand (predetermined) leaves per acre for Treatments 1, 2, and 3, respectively. Fertilizer and sucker control materials were applied approximately in proportion to the number of leaves. Hence, leaves per acre serve as an indicator of the range in treatment intensity. Other practices, including variety, were constant for all treatments within each experiment. All tobacco was grown, harvested, cured and prepared for sale in the conventional way.

A survey of previous unpublished research work indicated that using comparatively small experiment station plots, which ordinarily suffice for agronomic experiments, may result in inaccurate labor time

*Garnet L. Bradford is assistant professor of Agricultural Economics and Statistics, University of Kentucky.

measurements. Consequently, each treatment was applied to a minimum of two acres—the minimum acreage normally accomodated by a conventional flue-cured tobacco curing barn. Each treatment was applied to two separate plots of approximately one acre each, arranged in randomized complete block designs. Plots of this size were considered to be sufficiently large to obtain accurate labor time measurements. Specific details on treatment design, experimental design and measurement procedures are given by Bradford and Nelson [1].

A scatter diagram of priming labor plotted against harvested leaves per acre illustrated the type of measurements which were common for most of the 15 individual labor operations (Fig. 1). Three distinct characters (shown in the legend) are used to identify the three treatments. Locations are identified by letters and years by numbers. The 27 observations are treatment averages from each experiment.

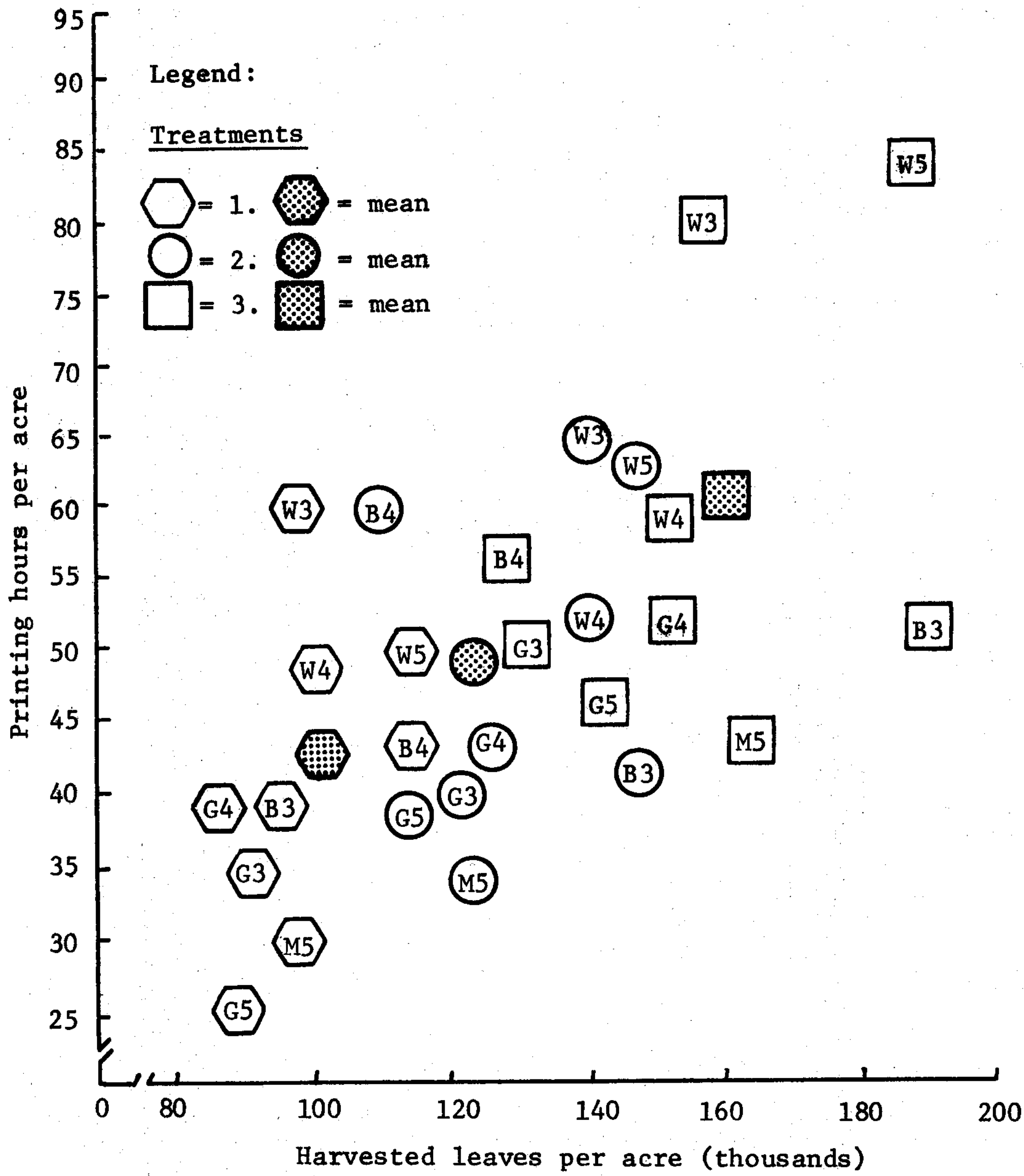

FIGURE 1. TREATMENT-AVERAGE OBSERVATIONS WITH PRIMING LABOR RELATED TO HARVESTED LEAVES PER ACRE[a]

[a]Locations are denoted by the following letters: W. Wayne County; B, Bertie County; G, Granville County; and M, Moore County. Years are denoted by the following digits: 3, 1963; 4, 1964; and 5, 1965.

PRIMING LABOR ANALYSIS

Procedures for deriving realistic estimates of the functional coefficients for priming labor involved making a series of statistical tests of the hypotheses previously enumerated. Initial regression models were selected on the basis of information contained in the study by Hunt, et al. [6], other similar studies, and inspection of scatter diagrams. Intermediate regression models were specified on the basis of F tests as applied to analysis of variance results. The choice of a final regression model involved using F and "t" tests to determine statistical significance after successively adding or deleting independent variables in alternative models.

Final results from the regression analysis of priming labor are summarized in Table 1. The eight dummy variables account for year and location-within-year variation. Harvested leaves per acre correspond to treatment variation. Interaction sources of variation (treatment X year and treatment X location-within-year) are not accounted for in this model. In other models, these sources of variation were specified by cross products of observations for production variables (harvested leaves, etc.) and year or location-within-year dummy variables. Hence, these latter variables are referred to as slope-changing dummy variables, because observations for each variable were either a zero or the counterpart continuous value for the corresponding production variable.

A variety of jargon has been used in defining and describing models containing these different types of variables. In subsequent discussion, the general term (regression models) will be employed, notwithstanding that such models often are given more precise or complicated sounding names. Economists have referred to such models by (1) mixed, (2) covariance estimation (3) dummy variable, (4) linear unspecified, and numerous other terms. A more thorough description of the properties of such models is given

TABLE 1. PRIMING LABOR, FINAL REGRESSION RESULTS

Item	Regression statistic		
	Parameter of coefficient	Estimate[a]	"t" value
General intercept value[b]	α	16.00	- - - - -[c]
Regression constant term	$\alpha + \alpha_3 + \gamma_1 + \gamma_1\alpha_3$	32.98	- - - - -[c]
Dummy variables (symbols):			
1964 (T_4)	$\alpha_4 - \alpha_3$	−11.79*	−3.53
1965 (T_5)	$\alpha_5 - \alpha_3$	− 6.46	−1.92
Bertie Co. (L_2)	$\gamma_2 - \gamma_1$	−26.64*	−7.93
Granville Co. (L_3)	$\gamma_3 - \gamma_1$	−21.84*	−6.48
Moore Co. (L_4)	$\gamma_4 - \gamma_1$	−22.48*	−6.61
Bertie Co., 1964 (L_2T_4)	$\gamma_2\alpha_4 - \gamma_1\alpha_3$	−27.46*	5.76
Granville Co., 1964 (L_3T_4)	$\gamma_3\alpha_4 - \gamma_1\alpha_3$	12.83*	2.71
Granville Co., 1965 (L_3T_5)	$\gamma_3\alpha_5 - \gamma_1\alpha_3$	1.83	.39
Harvested leaves (1,000)	β	.263	8.87

[a]In units of hours per acre, except for the estimate of β (.263), which is hours per 1,000 harvested leaves. Single asterisks indicate significance at the one percent level; the absence of asterisks indicates nonsignificance.

[b]Computed by multiplying the general mean value for harvested leaves per acre (129,000) by the slope regression coefficient (.263) and subtracting the resultant product from the general mean value for priming hours per acre (49.9)

[c]Not available.

by Johnston [7, pp. 221-228] or Graybill [3, pp. 383-403].

Regression coefficients for the dummy variables are estimates of linear combinations of parameters specified in the table. These combinations are a result of the reparameterization process which was used to avoid perfect multicollinearity. This process, commonly employed, eliminated specific discrete variables by combining parameters in the original ("nonreparameterized") form of the model. In the model, on which Table 1 results are based, zero-one variables representing 1963 (T_3) and the Wayne County location (L_1) were eliminated. Since estimates of the dummy-variable coefficients are invariant, different (desired) contrasts of these coefficients may be obtained by subtracting the coefficients for T_i and T_j, $i \neq j$, e.g., $\alpha_5 - \alpha_4 = (\alpha_5 - \alpha_3) - (\alpha_4 - \alpha_3)$, or coefficients for L_s and L_k, $s \neq k$.

Explanatory Variable Hypotheses

Testing these hypotheses involved selecting individual production variables and/or combinations of production variables for inclusion in alternative models. R^2 values were used to select the "best" explanatory variable in regressions where individual production variables were employed. F tests were used to test the significance of including two or more production variables.

The linear slope coefficient for priming labor (.263 hours per 1,000 harvested leaves) was highly significant. The R^2 value for this model (Table 1) was .86. When pounds per acre were used as an alternative to harvested leaves, the linear slope coefficient was estimated to be 2.096 hours per 100 pounds. But, the R^2 value corresponding to this alternative model dropped to .82. Since the error degrees of freedom were identical for both models, it was concluded that harvested leaves were more efficient estimates of changes in priming labor.

A third model included both pounds and leaves as continuous explanatory variables. But, this distorted estimates of the linear slope coefficients beyond reasonable interpretation. Obviously this was due to the high correlation between these two variables. Such correlation was expected since the experiments of this study were designed to obtain higher yields through use of more leaves and near-proportional increases in fertilizer amounts and other inputs per acre. In any event, use of models which included both pounds and leaves as continuous explanatory variables did not result in significant reductions in the error sum of squares.

Linearity Hypotheses

Testing these hypotheses involved adding quadratic forms of production variables to regression models which included only linear variables and then comparing estimates from both types of models. For priming labor, observations for the "leaves-linear" variable were entered as deviates from the general mean; this resulted in obersvations below 129 thousand leaves being negative and those above being positive. For the "leaves-squared" variable, all observations were squares of the linear deviates. Such a transformation procedure frequently is employed to reduce estimation bias resulting from intercorrelation of ordinary linear and squared terms.

Addition of a quadratic variable failed to reduce significantly the error sum of squares for each of the other 14 labor operations. For priming labor, the R^2 value increased only from .86 to .87; the "t" value for the quadratic coefficient was only -.46; the slope regression coefficient increased only from .263 to .274.

Intercept Hypotheses

Intercept values depend upon the independent variables in the model and the method of reparameterization. For example, one estimate of the priming labor intercept was 32.98 hours per acre (Table 1). This is an estimate of $\alpha + \alpha_3 + \gamma_1 + \gamma_1\alpha_3$ and would have been different had: (1) the set of dummy variables accounting for year variation been excluded and/or (2) reparameterization of location-within-year dummy variables been affected by deleting the Granville County location (L_3) rather than the Wayne County location (L_1).

What was desired, however, was an estimate of α alone. But, intercept values estimated using reparameterized models always contain unwanted effects. Consequently, an indirect estimation procedure was employed in order to make an alternative test of this hypothesis.

This procedure involved using the same observations, illustrated in Figure 1, to fit simple linear regression models. Specifically, simple linear regression models were fitted through the origin and then with intercept values. An F test was made to determine the significance of the intercept value, viz.,

$$F = \frac{\text{Additional reduction in the error sum of squares due to including the intercept term}}{\text{error mean square of the through-the-origin regression.}}$$

If this F value was significant and the simple regression intercept value was less (absolutely) than the intercept value obtained by the following equation:

$$(1)\ \hat{\alpha} = \bar{P} - \beta\bar{X}$$

where

$\overline{P}$ = the general mean for priming.

$\overline{X}$ = the general mean for harvested leaves, and

β = the regression coefficient estimated using the final model containing dummy variables (Table 1).

then it was concluded that the true intercept value (α) was significantly different from zero.

This procedure is weak in that it does not allow a decision on significance if the simple linear intercept value is greater than the value calculated using equation (1) above. However, it has the merit of lowering the probability of a Type I error, i.e., compared to using an F test involving only the simple linear intercept value.

Uniformity Hypotheses

Tests of these hypotheses involved adding or deleting zero-one or slope-changing dummy variables, refitting resultant models and then evaluating the significance of changes in the error sum of squares.

Adding various sets of slope-changing dummy variables did not significantly reduce the error sum of squares for any of the 15 labor operations—consistent with the lack of significant treatment X year or treatment X location in the ANOVA results reported by Bradford and Nelson [1]. For example, it was verified that the same priming labor slope value (.263 hours per 1,000 harvested leaves) applied to all years and locations. Slope-changing dummy variables were highly correlated with harvested leaves, so their addition biased the estimate of the slope value. However, this estimate (.263) was not changed significantly by deleting the two nonsignificant zero-one dummy variables (T_5 and L_3K_5). This was consistent with the low correlation of these variables with harvested leaves.

In contrast to slope values, intercept values were quite variable among years and locations. This is indicated by the scatter diagram (Figure 1) and demonstrated by the dummy-variable coefficients (Table 1). Large "t" values indicate that all except two of the dummy coefficients were highly significant. The exact differences shown, of course, vary with the reparameterization bases used in Table 1.

Stalk Position Hypotheses

ANOVA tests implied that priming cost per 1,000 harvested leaves did not vary significantly among stalk positions. This lack of significance was fairly uniform among treatments as was demonstrated by general nonsignificance of treatment X stalk position interaction variation. Thus, the slope coefficient shown in Table 1 (.263 per 1,000 harvested leaves) was hypothesized to apply to all four stalk positions. This hypothesis could not be rejected on the basis of "t" tests of differences between changes of slope coefficients. Coefficients varied from a high of .292 for the lower position of the leaves to a low of .227 for the mid-upper position of the leaves, but differences were not large enough to be judged statistically significant.

CONCLUSIONS AND IMPLICATIONS

In general, tests of the hypotheses indicate that: (1) each individual labor cost is linearly related to only one production variable, e.g., priming labor to harvested leaves, (2) slope coefficients are comparatively stable among different farm locations, years and stalk positions of the leaves, and (3) intercept coefficients may vary widely among locations and years. Such results suggest that a relatively simple procedure may be employed to estimate what labor costs might have been, had different production practices been used, viz., multiply the change in the quantity of the input by the slope coefficient and add (subtract) this product to the labor requirement for the higher (lower) level of the input. Suppose, for example, that labor requirements were 38 hours to prime 90 thousand leaves per acre in 1965 at Location 1. To have primed 140 thousand leaves, thus, would have required 38 + (.263) (50) = 51.2 hours per acre.

An obvious limitation of this procedure is its lack of strict validity when applied to future years and/or different farm locations. One is likely to be faced with an unknown but much lower (or higher) intercept or starting value for the labor operation; for example, priming labor for 90 thousand leaves may be 28 or 52 hours. In many cases, however, it would appear sufficient to make only some reasonable estimate of the change in the labor requirement, given a certain quantity change in the input. If so, slope coefficients derived in this study may be an improvement over the "typical" budget coefficients of the past.

REFERENCES

1. Bradford, G. L. and L. A. Nelson, *Labor Costs in Conventional Production of Flue-Cured Tobacco: Their Magnitude and Variability*, Tech. Bul. No. 190, Raleigh, North Carolina Agr. Exp. Sta., 1969.

2. Coutu, A. J. and F. A. Mangum, *Farm Management Manual.* Agr. Econ. Misc. Publ. No. 1, Raleigh, Dept. of Agr. Econ., North Carolina State College, 1960.

3. Graybill, F. A., *An Introduction to Linear Statistical Models,* New York, McGraw-Hill Book Company, Inc., 1961.

4. Greene, R. E. L., *Cost of Production Farm Products in North Carolina,* Bul. 305, Raleigh, North Carolina Agr. Exp. Sta., 1936.

5. Hartman, L. M. and G. S. Tolley, *Effects of Federal Acreage Controls on Costs and Techniques of Producing Flue-Cured Tobacco,* Tech. Bul. No. 146, Raleigh, North Carolina Agr. Exp. Sta., 1961.

6. Hunt, J. B., Jr., W. D. Toussaint and W. G. Woltz, *Acreage Controls and Poundage Controls: Their Effects on Most Profitable Production Practices for Flue-Cured Tobacco,* Tech. Bul. No. 162, Raleigh, North Carolina Agr. Exp. Sta., 1964.

7. Johnston, J. *Econometric Methods*, New York, McGraw-Hill Book Company, Inc. 1963.

8. North Carolina Agricultural Extension Service, *A Cost and Returns Guide for Selected Field Crops in North Carolina,* Extension Cir. No. 462, Raleigh, North Carolina State University, 1965.

9. Pierce, W. H. and M. S. Williams, *Cost of Producing Farm Products in North Carolina*, A. E. Inf. Ser. No. 29, Raleigh, Dept. of Agr. Econ., North Carolina State College, 1952.

10. Seagraves, J. A. and R. C. Manning, *Flue-Cured Tobacco Allotment Values and Uncertainty, 1934-1962,* Econ. Res. Rep. No. 2, Raleigh, Department of Economics, North Carolina State University, 1967.